WEST

Help | Logout | Interrupt

Main Menu | Search Form | Posting Counts | Show S Numbers | Edit S Numbers | Preferences | Cases

Search Results -

Terms	Documents
L6 and l3	4

Database: US Patents Full-Text Database
US Pre-Grant Publication Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search: L7

Refine Search

Recall Text | Clear

Search History

DATE: Monday, June 09, 2003 Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
DB=USPT; PLUR=YES; OP=OR			
L7	L6 and l3	4	L7
L6	L5 and l4	4	L6
L5	E222G	7	L5
L4	f64L	46	L4
L3	chromophore and l1	173	L3
L2	chromphore and L1	5	L2
L1	gfp	1681	L1

END OF SEARCH HISTORY

WEST

Generate Collection | Print

Search Results - Record(s) 1 through 4 of 4 returned.

☐ 1. Document ID: US 6403374 B1

L7: Entry 1 of 4 File: USPT Jun 11, 2002

US-PAT-NO: 6403374
DOCUMENT-IDENTIFIER: US 6403374 B1
**** See image for Certificate of Correction ****

TITLE: Long wavelength engineered fluorescent proteins

DATE-ISSUED: June 11, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Tsien; Roger Y.	La Jolla	CA		
Remington; S. James	Eugene	OR		
Cubitt; Andrew B.	San Diego	CA		
Heim; Roger	Del Mar	CA		
Ormo ; Mats F.	Huddinge			SE

US-CL-CURRENT: 435/325; 435/252.3, 435/252.33, 435/254.11, 435/320.1, 435/410, 536/23.1, 536/23.4, 536/23.6

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw Desc | Image

☐ 2. Document ID: US 6124128 A

L7: Entry 2 of 4 File: USPT Sep 26, 2000

US-PAT-NO: 6124128
DOCUMENT-IDENTIFIER: US 6124128 A

TITLE: Long wavelength engineered fluorescent proteins

DATE-ISSUED: September 26, 2000

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Tsien; Roger Y.	La Jolla	CA		
Cubitt; Andrew B.	San Diego	CA		
Heim; Roger	Del Mar	CA		
Ormo; Mats F.	Huddinge			SE
Remington; S. James	Eugene	OR		

US-CL-CURRENT: 435/252.33; 435/252.3, 435/320.1, 536/23.1, 536/23.5

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC
Draw Desc | Image

☐ 3. Document ID: US 6077707 A

L7: Entry 3 of 4 File: USPT Jun 20, 2000

US-PAT-NO: 6077707
DOCUMENT-IDENTIFIER: US 6077707 A

TITLE: Long wavelength engineered fluorescent proteins

DATE-ISSUED: June 20, 2000

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Tsien; Roger Y.	La Jolla	CA		
Remington; S. James	Eugene	OR		
Cubitt; Andrew B.	San Diego	CA		
Heim; Roger	Del Mar	CA		
Ormo; Mats F.	Huddinge			SE

US-CL-CURRENT: 435/325; 435/252.3, 435/252.33, 435/254.11, 435/320.1, 435/410, 435/69.1, 530/350, 536/23.1, 536/23.5

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | KWIC
Draw Desc | Image

☐ 4. Document ID: US 6054321 A

L7: Entry 4 of 4 File: USPT Apr 25, 2000

US-PAT-NO: 6054321
DOCUMENT-IDENTIFIER: US 6054321 A

TITLE: Long wavelength engineered fluorescent proteins

DATE-ISSUED: April 25, 2000

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Tsien; Roger Y.	La Jolla	CA		
Remington; S. James	Eugene	OR		
Cubitt; Andrew B.	San Diego	CA		
Heim; Roger	Del Mar	CA		
Ormo; Mats F.	Huddinge			SE

US-CL-CURRENT: 436/86; 530/350, 702/19, 702/22

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | KWIC
Draw Desc | Image

L7: Entry 1 of 4 File: USPT Jun 11, 2002

DOCUMENT-IDENTIFIER: US 6403374 B1
**** See image for Certificate of Correction ****
TITLE: Long wavelength engineered fluorescent proteins

Brief Summary Text (2):
Fluorescent molecules are attractive as reporter molecules in many assay systems because of their high sensitivity and ease of quantification. Recently, fluorescent proteins have been the focus of much attention because they can be produced in vivo by biological systems, and can be used to trace intracellular events without the need to be introduced into the cell through microinjection or permeabilization. The green fluorescent protein of Aequorea victoria is particularly interesting as a fluorescent protein. A cDNA for the protein has been cloned. (D. C. Prasher et al., "Primary structure of the Aequorea victoria green-fluorescent protein," Gene (1992) 111:229-33.) Not only can the primary amino acid sequence of the protein be expressed from the cDNA, but the expressed protein can fluoresce. This indicates that the protein can undergo the cyclization and oxidation believed to be necessary for fluorescence. Aequorea green fluorescent protein ("GFP") is a stable, proteolysis-resistant single chain of 238 residues and has two absorption maxima at around 395 and 475 nm. The relative amplitudes of these two peaks is sensitive to environmental factors (W. W. Ward. Bioluminescence and Chemiluminescence (M. A. DeLuca and W. D. McElroy, eds) Academic Press pp.235-242 (1981); W. W. Ward & S. H. Bokman Biochemistry 21:4535-4540 (1982); W. W. Ward et al. Photochem. Photobiol. 35:803-808 (1982)) and illumination history (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995)), presumably reflecting two or more ground states. Excitation at the primary absorption peak of 395 nm yields an emission maximum at 508 nm with a quantum yield of 0.72-0.85 (O. Shimomura and F. H. Johnson J. Cell. Comp. Physiol 59:223 (1962); J. G. Morin and J. W. Hastings, J. Cell. Physiol. 77:313 (1971); H. Morise et al. Biochemistry 13:2656 (1974); W. W. Ward Photochem. Photobiol. Reviews (Smith, K. C. ed.) 4:1 (1979); A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995); D. C. Prasher Trends Genet. 11:320-323 (1995); M. Chalfie Photochem. Photobiol. 62:651-656 (1995); W. W. Ward. Bioluminescence and Chemiluminescence (M. A. DeLuca and W. D. McElroy, eds) Academic Press pp. 235-242 (1981); W. W. Ward & S. H. Bokman Biochemistry 21:4535-4540 (1982); W. W. Ward et al. Photochem. Photobiol. 35:803-808 (1982)). The fluorophore results from the autocatalytic cyclization of the polypeptide backbone between residues Ser.sup.65 and Gly.sup.67 and oxidation of the .alpha.-.beta. bond of Tyr.sup.66 (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995); C. W. Cody et al. Biochemisty 32:1212-1218 (1993); R. Heim et al. Proc. Natl. Acad. Sci. USA 91:12501-12504 (1994)). Mutation of Ser.sup.65 to Thr (S65T) simplifies the excitation spectrum to a single peak at 488 nm of enhanced amplitude (R. Heim et al. Nature 373:664-665 (1995)), which no longer gives signs of conformational isomers (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995)).

Drawing Description Text (2):
FIGS. 1A-1B. (A) Schematic drawing of the backbone of GFP produced by Molscript (J. P. Kraulis, J. Appl. Cryst., 24:946 (1991)). The chromophore is shown as a ball and stick model. (B) Schematic drawing of the overall fold of GFP. Approximate residue numbers mark the beginning and ending of the secondary structure elements.

Drawing Description Text (3):
FIGS. 2A-2C. (A) Stereo drawing of the chromophore and residues in the immediate vicinity. Carbon atoms are drawn as open circles, oxygen is filled and nitrogen is

shaded. Solvent molecules are shown as isolated filled circles. (B) Portion of the final 2F.sub.o -F.sub.c electron density map contoured at 1.0-, showing the electron density surrounding the chromophore. (C) Schematic diagram showing the first and second spheres of coordination of the chromophore. Hydrogen bonds are shown as dashed lines and have the indicated lengths in .ANG.. Inset: proposed structure of the carbinolamine intermediate that is presumably formed during generation of the chromophore.

Detailed Description Text (3):
In one aspect, this invention provides a nucleic acid molecule comprising a nucleotide sequence encoding a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least an amino acid substitution located no more than about 0.5 nm from the chromophore of the engineered fluorescent protein, wherein the substitution alters the electronic environment of the chromophore, whereby the functional engineered fluorescent protein has a different fluorescent property than Aequorea green fluorescent protein.

Detailed Description Text (4):
In one aspect this invention provides a nucleic acid molecule comprising a nucleotide sequence encoding a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least a substitution at T203 and, in particular, T203X, wherein X is an aromatic amino acid selected from H, Y, W or F, said functional engineered fluorescent protein having a different fluorescent property than Aequorea green fluorescent protein. In one embodiment, the amino acid sequence further comprises a substitution at S65, wherein the substitution is selected from S65G, S65T, S65A, S65L, S65C, S65V and S65I. In another embodiment, the amino acid sequence differs by no more than the substitutions S65T/T203H; S65T/T203Y; S72A/F64L/S65G/T203Y; S65G/V68L/Q69K/S72A/T203Y; S72A/S65G/V68L/T203Y; S65G/S72A/T203Y; or S65G/S72A/T203W. In another embodiment, the amino acid sequence further comprises a substitution at Y66, wherein the substitution is selected from Y66H, Y66F, and Y66W. In another embodiment, the amino acid sequence further comprises a mutation from Table A. In another embodiment, the amino acid sequence further comprises a folding mutation. In another embodiment, the nucleotide sequence encoding the protein differs from the nucleotide sequence of SEQ ID NO:1 by the substitution of at least one codon by a preferred mammalian codon. In another embodiment, the nucleic acid molecule encodes a fusion protein wherein the fusion protein comprises a polypeptide of interest and the functional engineered fluorescent protein.

Detailed Description Text (5):
In another aspect, this invention provides a nucleic acid molecule comprising a nucleotide sequence encoding a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least an amino acid substitution at L42, V61, T62, V68, Q69, Q94, N121, Y145, H148, V150, F165, I167, Q183, N185, L220, E222 (not E222G), or V224, said functional engineered fluorescent protein having a different fluorescent property than Aequorea green fluorescent protein. In one embodiment, amino acid substitution is:

Detailed Description Text (24):
In another aspect, this invention provides a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least an amino acid substitution located no more than about 0.5 nm from the chromophore of the engineered fluorescent protein, wherein the substitution alters the electronic environment of the chromophore, whereby the functional engineered fluorescent protein has a different fluorescent property than Aequorea green fluorescent protein.

Detailed Description Text (25):
In another aspect, this invention provides a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid

sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least the amino acid substitution at T203, and in particular, T203X, wherein X is an aromatic amino acid selected from H, Y, W or F, said functional engineered fluorescent protein having a different fluorescent property than Aequorea green fluorescent protein. In one embodiment, the amino acid sequence further comprises a substitution at S65, wherein the substitution is selected from S65G, S65T, S65A, S65L, S65C, S65V and S65I. In another embodiment, the amino acid sequence differs by no more than the substitutions S65T/T203H; S65T/T203Y; S72A/F64L/S65G/T203Y; S72A/S65G/V68L/T203Y; S65G/V68L/Q69K/S72A/T203Y; S65G/S72A/T203Y; or S65G/S72A/T203W. In another embodiment, the amino acid sequence further comprises a substitution at Y66, wherein the substitution is selected from Y66H, Y66F, and Y66W. In another embodiment, the amino acid sequence further comprises a folding mutation. In another embodiment, the engineered fluorescent protein is part of a fusion protein wherein the fusion protein comprises a polypeptide of interest and the functional engineered fluorescent protein.

Detailed Description Text (31):
In another aspect, this invention provides a method for engineering a functional engineered fluorescent protein having a fluorescent property different than Aequorea green fluorescent protein, comprising substituting an amino acid that is located no more than 0.5 nm from any atom in the chromophore of an Aequorea-related green fluorescent protein with another amino acid; whereby the substitution alters a fluorescent property of the protein. In one embodiment, the amino acid substitution alters the electronic environment of the chromophore.

Detailed Description Text (58):
The term "nucleic acid probe" refers to a nucleic acid molecule that binds to a specific sequence or sub-sequence of another nucleic acid molecule. A probe is preferably a nucleic acid molecule that binds through complementary base pairing to the full sequence or to a sub-sequence of a target nucleic acid. It will be understood that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. Probes are preferably directly labelled as with isotopes, chromophores, lumiphores, chromogens, fluorescent proteins, or indirectly labelled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or sub-sequence.

Detailed Description Text (81):
The term "fluorescent property" refers to the molar extinction coefficient at an appropriate excitation wavelength, the fluorescence quantum efficiency, the shape of the excitation spectrum or emission spectrum, the excitation wavelength maximum and emission wavelength maximum, the ratio of excitation amplitudes at two different wavelengths, the ratio of emission amplitudes at two different wavelengths, the excited state lifetime, or the fluorescence anisotropy. A measurable difference in any one of these properties between wild-type Aequorea GFP and the mutant form is useful. A measurable difference can be determined by determining the amount of any quantitative fluorescent property, e.g., the amount of fluorescence at a particular wavelength, or the integral of fluorescence over the emission spectrum. Determining ratios of excitation amplitude or emission amplitude at two different wavelengths ("excitation amplitude ratioing" and "emission amplitude ratioing", respectively) are particularly advantageous because the ratioing process provides an internal reference and cancels out variations in the absolute brightness of the excitation source, the sensitivity of the detector, and light scattering or quenching by the sample.

Detailed Description Text (84):
As used herein, the term "fluorescent protein" refers to any protein capable of fluorescence when excited with appropriate electromagnetic radiation. This includes fluorescent proteins whose amino acid sequences are either naturally occurring or engineered (i.e., analogs or mutants). Many cnidarians use green fluorescent proteins ("GFPs") as energy-transfer acceptors in bioluminescence. A "green fluorescent protein," as used herein, is a protein that fluoresces green light. Similarly, "blue fluorescent proteins" fluoresce blue light and "red fluorescent proteins" fluoresce red light. GFPs have been isolated from the Pacific Northwest

jellyfish, Aequorea victoria, the sea pansy, Renilla reniformis, and Phialidium gregarium. W. W. Ward et al., Photochem. Photobiol., 35:803-808 (1982); L. D. Levine et al., Comp. Biochem. Physiol., 72B:77-85 (1982).

Detailed Description Text (85):
A variety of Aequorea-related fluorescent proteins having useful excitation and emission spectra have been engineered by modifying the amino acid sequence of a naturally occurring GFP from Aequorea Victoria. (D. C. Prasher et al., Gene, 111:229-233 (1992); R. Heim et al., Proc. Natl. Acad. Sci., USA, 91:12501-04 (1994); U.S. patent application Ser. No. 08/337,915, filed Nov. 10, 1994; International application PCT/US95/14692, filed Nov. 10, 1995.)

Detailed Description Text (87):
Aequorea-related fluorescent proteins include, for example and without limitation, wild-type (native) Aequorea victoria GFP (D. C. Prasher et al., "Primary structure of the Aequorea victoria green fluorescent protein," Gene, (1992) 111:229-33), whose nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) are presented in FIG. 3; allelic variants of this sequence, e.g., Q80R, which has the glutamine residue at position 80 substituted with arginine (M. Chalfie et al., Science, (1994) 263:802-805); those engineered Aequorea-related fluorescent proteins described herein, e.g., in Table A or Table F, variants that include one or more folding mutations and fragments of these proteins that are fluorescent, such as Aequorea green fluorescent protein from which the two amino-terminal amino acids have been removed. Several of these contain different aromatic amino acids within the central chromophore and fluoresce at a distinctly shorter wavelength than wild type species. For example, engineered proteins P4 and P4-3 contain (in addition to other mutations) the substitution Y66H, whereas W2 and W7 contain (in addition to other mutations) Y66W. Other mutations both close to the chromophore region of the protein and remote from it in primary sequence may affect the spectral properties of GFP and are listed in the first part of the table below.

Detailed Description Text (88):
Additional mutations in Aequorea-related fluorescent proteins, referred to as "folding mutations," improve the ability of fluorescent proteins to fold at higher temperatures, and to be more fluorescent when expressed in mammalian cells, but have little or no effect on the peak wavelengths of excitation and emission. It should be noted that these may be combined with mutations that influence the spectral properties of GFP to produce proteins with altered spectral and folding properties. Folding mutations include: F64L, V68L, S72A, and also T44A, F99S, Y145F, N146I, M153T or A, V163A, I167T, S175G, S205T and N212K.

Detailed Description Text (94):
Fluorescent characteristics of Aequorea-related fluorescent proteins depend, in part, on the electronic environment of the chromophore. In general, amino acids that are within about 0.5 nm of the chromophore influence the electronic environment of the chromophore. Therefore, substitution of such amino acids can produce fluorescent proteins with altered fluorescent characteristics. In the excited state, electron density tends to shift from the phenolate towards the carbonyl end of the chromophore. Therefore, placement of increasing positive charge near the carbonyl end of the chromophore tends to decrease the energy of the excited state and cause a red-shift in the absorbance and emission wavelength maximum of the protein. Decreasing positive charge near the carbonyl end of the chromophore tends to have the opposte effect, causing a blue-shift in the protein's wavelengths.

Detailed Description Text (95):
Amino acids with charged (ionized D, E, K, and R), dipolar (H, N, Q, S, T, and uncharged D, E and K), and polarizable side groups (e.g., C, F, H, M, W and Y) are useful for altering the electronic environment of the chromophore, especially when they substitute an amino acid with an uncharged, nonpolar or non-polarizable side chain. In general, amino acids with polarizable side groups alter the electronic environment least, and, consequently, are expected to cause a comparatively smaller change in a fluorescent property. Amino acids with charged side groups alter the environment most, and, consequently, are expected to cause a comparatively larger change in a fluorescent property. However, amino acids with charged side groups are more likely to disrupt the structure of the protein and to prevent proper folding if

buried next to the chromophore without any additional solvation or salt bridging. Therefore charged amino acids are most likely to be tolerated and to give useful effects when they replace other charged or highly polar amino acids that are already solvated or involved in salt bridges. In certain cases, where substitution with a polarizable amino acid is chosen, the structure of the protein may make selection of a larger amino acid, e.g., W, less appropriate. Alternatively, positions occupied by amino acids with charged or polar side groups that are unfavorably oriented may be substituted with amino acids that have less charged or polar side groups. In another alternative, an amino acid whose side group has a dipole oriented in one direction in the protein can be substituted with an amino acid having a dipole oriented in a different direction.

Detailed Description Text (96):
More particularly, Table B lists several amino acids located within about 0.5 nm from the chromophore whose substitution can result in altered fluorescent characteristics. The table indicates, underlined, preferred amino acid substitutions at the indicated location to alter a fluorescent characteristic of the protein. In order to introduce such substitutions, the table also provides codons for primers used in site directed mutagenesis involving amplification. These primers have been selected to encode economically the preferred amino acids, but they encode other amino acids as well, as indicated, or even a stop codon, denoted by Z. In introducing substitutions using such degenerate primers the most efficient strategy is to screen the collection to identify mutants with the desired properties and then sequence their DNA to find out which of the possible substitutions is responsible. Codons are shown in double-stranded form with sense strand above, antisense stand below. In nucleic acid sequences, R=(A or g); Y=(C or T); M=(A or C); K=(g or T); S=(g or C); W=(A or T); H=(A, T, or C); B=(g, T, or C); V=(g, A, or C); D=(g, A, or T); N=(A, C, g, or T).

Detailed Description Text (98):
In another embodiment, an amino acid that is close to a second amino acid within about 0.5 nm of the chromophore can, upon substitution, alter the electronic properties of the second amino acid, in turn altering the electronic environment of the chromphore. Table D presents two such amino acids. The amino acids, L220 and V224, are close to E222 and oriented in the same direction in the .beta. pleated sheet.

Detailed Description Text (100):
One embodiment of the invention includes a nucleic acid molecule comprising a nucleotide sequence encoding a functional engineered fluorescent protein whose amino acid sequence is substantially identical to the amino acid sequence of Aequorea green fluorescent protein (SEQ ID NO:2) and which differs from SEQ ID NO:2 by at least a substitution at E222, but not including E222G, wherein the functional engineered fluorescent protein has a different fluorescent property than Aequorea green fluorescent protein. Preferably, the substitution at E222 is selected from the group of N and Q. The E222 substitution can be combined with other mutations to improve the properties of the protein, such as a functional mutation at F64.

Detailed Description Text (152):
As a step in understanding the properties of GFP, and to aid in the tailoring of GFPs with altered characteristics, we have determined the three dimensional structure at 1.9 .ANG. resolution of the S65T mutant (R. Heim et al. Nature 373:664-665 (1995)) of A. victoria GFP. This mutant also contains the ubiquitous Q80R substitution, which accidentally occurred in the early distribution of the GFP cDNA and is not known to have any effect on the protein properties (M. Chalfie et al. Science 263:802-805 (1994)).

Detailed Description Text (153):
Histidine-tagged S65T GFP (R. Heim et al. Nature 373:664-665 (1995)) was overexpressed in JM109/pRSET.sub.B in 4 1 YT broth plus ampicillin at 37.degree., 450 rpm and 5 l/min air flow. The temperature was reduced to 25.degree. at A.sub.595 =0.3, followed by induction with 1 mM isopropylthiogalactoside for 5 h. Cell paste was stored at -80.degree. overnight, then was resuspended in 50 mM HEPES pH 7.9, 0.3 M NaCl, 5 mM 2-mercaptoethanol, 0.1 mM phenylmethyl-sulfonylfluoride (PMSF), passed once through a French press at 10,000 psi, then centrifuged at 20 K rpm for 45 min.

The supernatant was applied to a Ni-NTA-agarose column (Qiagen), followed by a wash with 20 mM imidazole, then eluted with 100 mM imidazole. Green fractions were pooled and subjected to chymotryptic (Sigma) proteolysis (1:50 w/w) for 22 h at RT. After addition of 0.5 mM PMSF, the digest was reapplied to the Ni column. N-terminal sequencing verified the presence of the correct N-terminal methionine. After dialysis against 20 mM HEPES, pH 7.5 and concentration to A.sub.490 =20, rod-shaped crystals were obtained at RT in hanging drops containing 5 .mu.L protein and 5 .mu.L well solution, 22-26% PEG 4000 (Serva), 50 mM HEPES pH 8.0-8.5, 50 mM MgCl.sub.2 and 10 mM 2-mercapto-ethanol within 5 days. Crystals were 0.05 mm across and up to 1.0 mm long. The space group is P2.sub.1 2.sub.1 2.sub.1 with a=51.8, b=62.8, c=70.7 .ANG., Z=4. Two crystal forms of wild-type GFP, unrelated to the present form, have been described by M. A. Perrozo, K. B. Ward, R. B. Thompson, & W. W. Ward. J. Biol. Chem. 203, 7713-7716 (1988).

Detailed Description Text (154):
The structure of GFP was determined by multiple isomorphous replacement and anomalous scattering (Table E), solvent flattening, phase combination and crystallographic refinement. The most remarkable feature of the fold of GFP is an eleven stranded .beta.-barrel wrapped around a single central helix (FIGS. 1A and 1B), where each strand consists of approximately 9-13 residues. The barrel forms a nearly perfect cylinder 42 .ANG. long and 24 .ANG. in diameter. The N-terminal half of the polypeptide comprises three anti-parallel strands, the central helix, and then 3 more anti-parallel strands, the latter of which (residues 118-123) is parallel to the N-terminal strand (residues 11-23). The polypeptide backbone then crosses the "bottom" of the molecule to form the second half of the barrel in a five-strand Greek Key motif. The top end of the cylinder is capped by three short, distorted helical segments, while one short, very distorted helical segment caps the bottom of the cylinder. The main-chain hydrogen bonding lacing the surface of the cylinder very likely accounts for the unusual stability of the protein towards denaturation and proteolysis. There are no large segments of the polypeptide that could be excised while preserving the intactness of the shell around the chromophore. Thus it would seem difficult to re-engineer GFP to reduce its molecular weight (J. Dopf & T. M. Horiagon Gene 173:39-43 (1996)) by a large percentage.

Detailed Description Text (155):
The p-hydroxybenzylideneimidazolidinone chromophore (C. W. Cody et al. Biochemistry 32:1212-1218 (1993)) is completely protected from bulk solvent and centrally located in the molecule. The total and presumably rigid encapsulation is probably responsible for the small Stokes' shift (i.e. wavelength difference between excitation and emission maxima), high quantum yield of fluorescence, inability of O.sub.2 to quench the excited state (B. D. Nageswara Rao et al. Biophys. J. 32:630-632 (1980)), and resistance of the chromophore to titration of the external pH (W. W. Ward. Bioluminescence and Chemiluminescence (M. A. DeLuca and W. D. McElroy, eds) Academic Press pp. 235-242 (1981); W. W. Ward & S. H. Bokman. Biochemistry 21:4535-4540 (1982); W. W. Ward et al. Photochem. Photobiol. 35:803-808 (1982)). It also allows one to rationalize why fluorophore formation should be a spontaneous intramolecular process (R. Heim et al. Proc. Nat. Acad. Sci. USA 91:12501-12504 (1994)), as it is difficult to imagine how an enzyme could gain access to the substrate. The plane of the chromophore is roughly perpendicular (60.degree.) to the symmetry axis of the surrounding barrel. One side of the chromophore faces a surprisingly large cavity, that occupies a volume of approximately 135 .ANG..sup.3 (B. Lee & F. M. Richards. J. Mol. Biol. 55:379-400 (1971)). The atomic radii were those of Lee & Richards, calculated using the program MS with a probe radius of 1.4 .ANG.. (M. L. Connolly, Science 221:709-713 (1983)). The cavity does not open out to bulk solvent. Four water molecules are located in the cavity, forming a chain of hydrogen bonds linking the buried side chains of Glu.sup.222 and Gln.sup.69. Unless occupied, such a large cavity would be expected to de-stabilize the protein by several kcal/mol (S. J. Hubbard et al., Protein Engineering 7:613-626 (1994); A. E. Eriksson et al. Science 255:178-183 (1992)). Part of the volume of the cavity might be the consequence of the compaction resulting from cyclization and dehydration reactions. The cavity might also temporarily accommodate the oxidant, most likely O.sub.2 (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995); R. Heim et al. Proc. Natl. Acad. Sci. USA 91:12501-12504 (1994); S. Inouye & F. I. Tsuji. FEBS Lett. 351:211-214 (1994)), that dehydrogenates the .alpha.-.beta.-bond of Tyr.sup.66. The chromophore, cavity, and

side chains that contact the chromophore are shown in FIG. 2A and a portion of the final electron density map in this vicinity in 2B.

Detailed Description Text (156):
The opposite side of the chromophore is packed against several aromatic and polar side chains. Of particular interest is the intricate network of polar interactions with the chromophore (FIG. 2C). His.sup.148, Thr.sup.203 and Ser.sup.203 form hydrogen bonds with the phenolic hydroxyl; Arg.sup.96 and Gin.sup.94 interact with the carbonyl of the imidazolidinone ring and Glu.sup.222 forms a hydrogen bond with the side chain of Thr.sup.65. Additional polar interactions, such as hydrogen bonds to Arg.sup.96 from the carbonyl of Thr.sup.62, and the side-chain carbonyl of Gln.sup.183, presumably stabilize the buried Arg.sup.96 in its protonated form. In turn, this buried charge suggests that a partial negative charge resides on the carbonyl oxygen of the imidazolidinone ring of the deprotonated fluorophore, as has previously been suggested (W. W. Ward. Bioluminescence and Chemiluminescence (M. A. DeLuca and W. D. McElroy, eds) Academic Press pp. 235-242 (1981); W. W. Ward & S. H. Bokman. Biochemistry 21:4535-4540 (1982); W. W. Ward et al. Photochem. Photobiol. 35:803-808 (1982)). Arg.sup.96 is likely to be essential for the formation of the fluorophore, and may help catalyze the initial ring closure. Finally, Tyr.sup.145 shows a typical stabilizing edge-face interaction with the benzyl ring. Trp.sup.57, the only tryptophan in GFP, is located 13 .ANG. to 15 .ANG. from the chromophore and the long axes of the two ring systems are nearly parallel. This indicates that efficient energy transfer to the latter should occur, and explains why no separate tryptophan emission is observable (D. C. Prasher et al. Gene 111:229-233 (1992). The two cysteines in GFP, Cys.sup.48 and Cys.sup.70, are 24 .ANG. apart, too distant to form a disulfide bridge. Cys.sup.70 is buried, but Cys.sup.48 should be relatively accessible to sulfhydryl-specific reagents. Such a reagent, 5,5'-dithiobis(2-nitrobenzoic acid), is reported to label GFP and quench its fluorescence (S. Inouye & F. I. Tsuji FEBS Lett. 351:211-214 (1994)). This effect was attributed to the necessity for a free sulfhydryl, but could also reflect specific quenching by the 5-thio-2-nitrobenzoate moiety that would be attached to Cys.sup.48.

Detailed Description Text (157):
Although the electron density map is for the most part consistent with the proposed structure of the chromophore (D. C. Prasher et al. Gene 111:229-233 (1992); C. W. Cody et al. Biochemistry 32:1212-1218 (1993)) in the cis [Z-] configuration, with no evidence for any substantial fraction of the opposite isomer around the chromophore double bond, difference features are found at >.sigma. in the final (F.sub.o -F.sub.c) electron density map that can be interpreted to represent either the intact, uncyclized polypeptide or a carbinolamine (inset to FIG. 2C). This suggests that a significant fraction, perhaps as much as 30% of the molecules in the crystal, have failed to undergo the final dehydration reaction. Confirmation of incomplete dehydration comes from electrospray mass spectrometry, which consistently shows that the average masses of both wild-type and S65T GFP (31,086.+-.4 and 31,099.5.+-.4 Da, respectively) are 6-7 Da higher than predicted (31,079 and 31,093 Da, respectively) for the fully matured proteins. Such a discrepancy could be explained by a 30-35% mole fraction of apoprotein or carbinolamine with 18 or 20 Da higher molecular weight The natural abundance of .sup.13 C and .sup.2 H and the finite resolution of the Hewlett-Packard 5989B electrospray mass spectrometer used to make these measurements do not permit the individual peaks to be resolved, but instead yields an average mass peak with a full width at half maximum of approximately 15 Da. The molecular weights shown include the His-tag, which has the sequence MRGSHHHHHH GMASMTGGQQM GRDLYDDDDK DPPAEF (SEQ ID NO:5). Mutants of GFP that increase the efficiency of fluorophore maturation might yield somewhat brighter preparations. In a model for the apoprotein, the Thr.sup.65 -Tyr.sup.66 peptide bond is approximately in the .alpha.-helical conformation, while the peptide of Tyre.sup.66 -Gly.sup.67 appears to be tipped almost perpendicular to the helix axis by its interaction with Arg.sup.96. This further supports the speculation that Arg.sup.96 is important in generating the conformation required for cyclization, and possibly also for promoting the attack of Gly.sup.67 on the carbonyl carbon of Thr.sup.65 (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995)).

Detailed Description Text (158):
The results of previous random mutagenesis have implicated several amino acid side

chains to have substantial effects on the spectra and the atomic model confirms that these residues are close to the chromophore. The mutations T203I and E222G have profound but opposite consequences on the absorption spectrum (T. Ehrig et al. FEBS Letters 367:163-166 (1995)). T203I (with wild-type Ser.sup.65) lacks the 475 nm absorbance peak usually attributed to the anionic chromophore and shows only the 395 nm peak thought to reflect the neutral chromophore (R. Heim et al. Proc. Natl. Acad. Sci. USA 91:12501-12504 (1994); T. Ehrig et al. FEBS Letters 367:163-166 (1995)). Indeed, Thr.sup.203 is hydrogen-bonded to the phenolic oxygen of the chromophore, so replacement by Ile should hinder ionization of the phenolic oxygen. Mutation of Glu.sup.222 to Gly (T. Ehrig et al. FEBS Letters 367:163-166 (1995)) has much the same spectroscopic effect as replacing Ser.sup.65 by Gly, Ala, Cys, Val, or Thr, namely to suppress the 395 nm peak in favor of a peak at 470-490 nm (R. Heim et al. Natutre 373:664-665 (1995); S. Delagrave et al. Bio/Technology 13:151-154 (1995)). Indeed Glu.sup.222 and the remnant of Thr.sup.65 are hydrogen-bonded to each other in the present structure, probably with the uncharged carboxyl of Glu.sup.222 acting as donor to the side chain oxygen of Thr.sup.65. Mutations E222G, S65G, S65A, and S65V would all suppress such H-bonding. To explain why only wild-type protein has both excitation peaks, Ser.sup.65, unlike Thr.sup.65, may adopt a conformation in which its hydroxyl donates a hydrogen bond to and stabilizes Glu.sup.222 as an anion, whose charge then inhibits ionization of the chromophore. The structure also explains why some mutations seem neutral. For example, Gln.sup.80 is a surface residue far removed from the chromophore, which explains why its accidental and ubiquitous mutation to Arg seems to have no obvious intramolecular spectroscopic effect (M. Chalfie et al. Science 263:802-805 (1994)).

Detailed Description Text (159):
The development of GFP mutants with red-shifted excitation and emission maxima is an interesting challenge in protein engineering (A. B. Cubitt et al. Trends Biochem. Sci. 20:448-455 (1995); R. Heim et al. Nature 373:664-665 (1995); S. Delagrave et al. Bio/Technology 13:151-154 (1995)). Such mutants would also be valuable for avoidance of cellular autofluorescence at short wavelengths, for simultaneous multicolor reporting of the activity of two or more cellular processes, and for exploitation of fluorescence resonance energy transfer as a signal of protein-protein interaction (R. Heim & R. Y. Tsien. Current Biol. 6:178-182 (1996)). Extensive attempts using random mutagenesis have shifted the emission maximum by at most 6 nm to longer wavelengths, to 514 nm (R. Heim & R. Y. Tsien. Current Biol. 6:178-182 (1996)); previously described "red-shifted" mutants merely suppressed the 395 nm excitation peak in favor of the 475 nm peak without anv sigiuficant reddening of the 505 nm emission (S. Delagrave et al. Bio/Technology 13:151-154 (1995)). Because Thr.sup.203 is revealed to be adjacent to the phenolic end of the chromophore, we mutated it to polar aromatic residues such as His, Tyr, and Trp in the hope that the additional polarizability of their .pi. systems would lower the energy of the excited state of the adjacent chromophore. All three substitutions did indeed shift the emission peak to greater than 520 nm (Table F). A particularly attractive mutation was T203Y/S65G/V68L/S72A, with excitation and emission peaks at 513 and 527 nm respectively. These wavelengths are sufficiently different from previous GFP mutants to be readily distinguishable by appropriate filter sets on a fluorescence microscope. The extinction coefficient, 36,500 M.sup.-1 cm.sup.-1, and quantum yield, 0.63, are almost as high as those of S65T (R. Heim et al. Nature 373:664-665 (1995)).

Detailed Description Text (160):
Comparison of Aequorea GFP with other protein pigments is instructive. Unfortunately, its closest characterized homolog, the GFP from the sea pansy Renilla reniformis (O. Shimomura and F. H. Johnson J. Cell. Comp. Physiol. 59:223 (1962); J. G. Morin and J. W. Hastings, J. Cell. Physiol. 77:313 (1971); H. Morise et al. Biochemistry 13:2656 (1974); W. W. Ward Photochem. Photobiol. Reviews (Smith, K. C. ed.) 4:1 (1979); W. W. Ward. Bioluminescence and Chemiluminescence (M. A. DeLuca and W. D. McElroy, eds) Academic Press pp. 235-242 (1981); W. W. Ward & S. H. Bokman Biochemistry 21:4535-4540 (1982); W. W. Ward et al. Photochem. Photobiol. 35:803-808 (1982)), has not been sequenced or cloned, though its chromophore is derived from the same FSYG sequence as in wild-type Aequorea GFP (R. M. San Pietro et al. Photochem. Photobiol. 57:63S (1993)). The closest analog for which a three dimensional structure is available is the photoactive yellow protein (PYP, G. E. O. Borgstahl et al. Biochemistry 34:6278-6287 (1995)), a 14-kDa photoreceptor from

halophilic bacteria. PYP in its native dark state absorbs maximally at 446 nm and transduces light with a quantum yield of 0.64, rather closely matching wild-type GFP's long wavelength absorbance maximum near 475 nm and fluorescence quantum yield of 0.72-0.85. The fundamental chromophore in both proteins is an anionic p-hydroxycinnamyl group, which is covalently attached to the protein via a thioester linkage in PYP and a heterocyclic iminolactam in GFP. Both proteins stabilize the negative charge on the chromophore with the help of buried cationic arginine and neutral glutamic acid groups, Arg.sup.52 and Glu.sup.46 in PYP and Arg.sup.96 and Glu.sup.222 in GFP, though in PYP the residues are close to the oxyphenyl ring whereas in GFP they are nearer the carbonyl end of the chromophore. However, PYP has an overall .alpha./.beta. fold with appropriate flexibility and signal transduction domains to enable it to mediate the cellular phototactic response, whereas GFP is a much more regular and rigid .beta.-barrel to minimize parasitic dissipation of the excited state energy as thermal or conformational motions. GFP is an elegant example of how a visually appealing and extremely useful function, efficient fluorescence, can be spontaneously generated from a cohesive and economical protein structure.

Detailed Description Text (161):
A. Summary of GFP Structure Determination

Detailed Description Text (162):
Data were collected at room temperature in house using either Molecular Structure Corp. R-axis II or San Diego Multiwire Systems (SDMS) detectors (Cu K.quadrature.) and later at beamline X4A at the Brookhaven National Laboratory at the selenium absorption edge (.quadrature.=0.979 .ANG.) using image plates. Data were evaluated using the HKL package (Z. Otwinowski, in Proceedings of the CCP4 Study Weekend: Data Collection and Processing, L. Sawyer, N. Issacs, S. Bailey, Eds. (Science and Engineering Research Council (SERC), Daresbury Laboratory, Warrington, UK, (1991)), pp 56-62; W. Minor, XDISPLAYF (Purdue University, West Lafayette, Ind., 1993)) or the SDMS software (A. J. Howard et al. Meth. Enzymol. 114:452-471 (1985)). Each data set was collected from a single crystal. Heavy atom soaks were 2 mM in mother liquor for 2 days. Initial electron density maps were based on three heavy atom derivatives using in-house data, then later were replaced with the synchrotron data. The EMTS difference Patterson map was solved by inspection, then used to calculate difference Fourier maps of the other derivatives. Lack of closure refinement of the heavy atom parameters was performed using the Protein package (W. Steigemann, in Ph.D. Thesis (Technical University, Munich, 1974)). The MIR maps were much poorer than the overall figure of merit would suggest, and it was clear that the EMTS isomorphous differences dominated the phasing. The enhanced anomalous occupancy for the synchrotron data provided a partial solution to the problem. Note that the phasing power was reduced for the synchrotron data, but the figure of merit was unchanged. All experimental electron density maps were improved by solvent flattening using the program DM of the CCP4 (CCP4: A Suite of Programs for Protein Crystallography (SERC Daresbury Laboratory, Warrington WA4 4AD UK, 1979)) package assuming a solvent content of 38%. Phase combination was performed with PHASCO2 of the Protein package using a weight of 1.0 on the atomic model. Heavy atom parameters were subsequently improved by refinement against combined phases. Model building proceeded with FRODO and O (T. A. Jones et al. Acta. Crystallogr. Sect. A 47:110 (1991); T. A. Jones, in Computational Crystallography D. Sayre, Ed. (Oxford University Press, Oxford, 1982) pp. 303-317) and crystallographic refinement was performed with the TNT package (D. E. Tronrud et al. Acta Cryst. A 43:489-503 (1987)). Bond lengths and angles for the chromophore were estimated using CHEM3D (Cambridge Scientific Computing). Final refinement and model building was performed against the X4A selenomethione data set, using (2F.sub.o -F.sub.c) electron density maps. The data beyond 1.9 .ANG. resolution have not been used at this stage. The final model contains residues 2-229 as the terminal residues are not visible in the electron density map, and the side chains of several disordered surface residues have been omitted. Density is weak for residues 156-158 and coordinates for these residues are unreliable. This disordering is consistent with previous analyses showing that residues 1 and 233-238 are dispensible but that further truncations may prevent fluorescence (J. Dopf & T. M. Horiagon. Gene 173:39-43 (1996)). The atomic model has been deposited in the Protein Data Bank (access code 1EMA).

Detailed Description Text (164):
The mutations F64L, V68L and S72A improve the folding of GFP at 37.quadrature. (B.

P. Cormack et al. Gene 173:33 (1996)) but do not significantly shift the emssion spectra.

Detailed Description Paragraph Table (3):
TABLE B Original position and presumed role Change to Codon L42 Aliphatic residue near C.dbd.N of CFHLQRWYZ 5'YDS 3' chromophore 3'RHS 5' V61 Aliphatic residue near central FYHCLR YDC --CH.dbd. of chromophore RHg T62 Almost directly above center of AVFS KYC chromophore bridge MRg DEHKNQ VAS BTS FYHCLR YDC RHg V68 Aliphatic residue near carbonyl FYHL YWC and G67 RWg N121 Near C-N Site of ring closure CFHLQRWYZ YDS between T65 and G67 RHS Y145 Packs near tyrosine ring of WCFL TKS chromophore AMS DEHNKQ VAS BTS H148 H-bonds to phenolate oxygen FYNI WWC WWg KQR MRg KYC V150 Aliphatic residue near tyrosine ring FYHL YWC of chromophore RWg F165 Packs near tyrosine ring CHQRWYZ YRS RYS I167 Aliphatic residue near phenolate; FYHL YWC I167T has effects RWg T203 H-bonds to phenolic oxygen of FHLQRWYZ YDS chromophore RHS E222 Protonation regulates ionization of HKNQ MAS chromophore KTS

Detailed Description Paragraph Table (4):
TABLE C Original position and presumed role Change to Codon Q69 Terminates chain of H-bonding waters KREG RRg YYC Q94 H-bonds to carbonyl terminus of DEHKNQ VAS chromophore BTS Q183 Bridges Arg96 and center of chromophore HY YAC bridge RTG EK RAg YTC N185 Part of H-bond network near carbonyl of DEHNKQ VAS chromophore BTS

Detailed Description Paragraph Table (5):
TABLE D Original position and presumed role Change to Codon L220 Packs next to Glu222; HKNPQT MMS to make GFP pH sensitive KKS Y224 Packs next to Glu222; HKNPQT MMS to make GFP pH sensitive KKS CFHLQRWYZ YDS RHS

Detailed Description Paragraph Table (7):
TABLE F Excita- Emis- tion Extinction sion max. coefficient max. Clone Mutations (nm) (10.sup.3 M.sup.-1 cm.sup.-1) (nm) S65T S65T 489 39.2 511 5B T203H/S65T 512 19.4 524 6C T203Y/S65T 513 14.5 525 10B T203Y/F64L/S65G/S72A 513 30.8 525 10C T203Y/F65G/V68L/S72A 513 36.5 527 11 T203W/S65G/S72A 502 33.0 512 12H T203Y/S65G/S72A 513 36.5 527 20A T203Y/S65G/V68L/Q69K/S72A 515 46.0 527

CLAIMS:

2. The nucleic acid molecule of claim 1 wherein the amino acid sequence differs by no more than the substitutions S65T/T203H; S65T/T203Y; S72A/F64L/S65G/T203Y; S72A/S65G/V68L/T203Y; S65G/V68L/Q69K/S72A/T203Y; S65G/S72A/T203Y; or S65G/S72A/T203W.

5. Tie nucleic acid molecule of claim 1 wherein the amino acid sequence further comprises a folding mutation selected from the group consisting of F64L, V68L and S72A.

9. The expression vector of claim 8 wherein the amino acid sequence differs by no more than the substitutions S65T/T203H; S65T/T203Y; S72A/F64L/S65G/T203Y; S72A/S65G/V68L/T203Y; S65G/V68L/Q69K/S72A/T203Y, S65G/S72A/T203Y; or S65G/S72A/T203W.

12. The expression vector of claim 8 wherein the amino acid sequence further comprises a folding mutation selected from the group consisting of F64L, V68L and S72A.

16. The recombinant host cell of claim 15 wherein the amino acid sequence differs by no more than the substitutions S65T/T203H; S65T/T203Y; S72A/F64L/S65G/T203Y; S72A/S65G/V68L/T203Y; S65G/V68L/Q69K/S72A/T203Y; S65G/S72A/T203Y; or S65G/S72A/T203W.

19. The recombinant host cell of claim 15 wherein the amino acid sequence further comprises a folding mutation selected from the group consisting of F64L, V68L and S72A.